BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

34.WO1

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, C07K 14/47, C12N 15/10, 15/66, C12Q 1/68, G01N 33/50, C07K 16/18, G01N 33/53, A61K 48/00, 38/17

A2

(11) International Publication Number: WO 99/06549

(43) International Publication Date: 11 February 1999 (11.02.99)

(21) International Application Number: PCT/IB98/01231

(22) International Filing Date: 31 July 1998 (31.07.98)

(30) Priority Data:
08/905,279 1 August 1997 (01.08.97) US

(71) Applicant *(for all designated States except US)*: GENSET [FR/FR]; 24, rue Royale, F–75008 Paris (FR).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: DUMAS MILNE EDWARDS, Jean–Baptiste [FR/FR]; 8, rue Grégoire de Tours, F–75006 Paris (FR). DUCLERT, Aymeric [FR/FR]; 6 ter, rue Victorine, F–94100 Saint–Maur (FR). LACROIX, Bruno [FR/FR]; 93, route de Vourles, F–69230 Saint–Genis Laval (FR).

(74) Agents: MARTIN, Jean–Jacques et al.; Cabinet Regimbeau, 26, avenue Kléber, F–75116 Paris (FR).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) **Title:** 5' ESTs FOR SECRETED PROTEINS EXPRESSED IN TESTIS AND OTHER TISSUES

(57) **Abstract**

The sequences of 5' ESTs derived from mRNAs encoding secreted proteins are disclosed. The 5' ESTs may be to obtain cDNAs and genomic DNAs corresponding to the 5' ESTs. The 5' ESTs may also be used in diganostic, forensic, gene therapy, and chromosome mapping procedures. Upstream regulatory sequences may also be obtained using the 5' ESTs. The 5' ESTs may also be used to design expression vectors and secretion vectors.

controlling a variety of human conditions. The 5' ESTs may also be used to obtain the corresponding genomic DNA. The term "corresponding genomic DNA" refers to the genomic DNA which encodes the mRNA from which the 5' EST was derived.

Alternatively, the 5' ESTs may be used to obtain and express extended cDNAs encoding portions of the secreted protein. The portions may comprise the signal peptides of the secreted proteins or the mature proteins generated when the signal peptide is cleaved off. The portions may also comprise polypeptides having at least 10 consecutive amino acids encoded by the extended cDNAs or full length cDNAs. Alternatively, the portions may comprise at least 15 consecutive amino acids encoded by the extended cDNAs or full length cDNAs. In some embodiments, the portions may comprise at least 25 consecutive amino acids encoded by the extended cDNAs or full length cDNAs. In other embodiments, the portions may comprise at least 40 amino acids encoded by the extended cDNAs or full length cDNAs.

Antibodies which specifically recognize the entire secreted proteins encoded by the extended cDNAs, full length cDNAs, or fragments thereof having at least 10 consecutive amino acids, at least 15 consecutive amino acids, at least 25 consecutive amino acids, or at least 40 consecutive amino acids may also be obtained as described below. Antibodies which specifically recognize the mature protein generated when the signal peptide is cleaved may also be obtained as described below. Similarly, antibodies which specifically recognize the signal peptides encoded by the extended cDNAs or full length cDNAs may also be obtained.

In some embodiments, the extended cDNAs obtained using the 5' ESTs include the signal sequence. In other embodiments, the extended cDNAs obtained using the 5' ESTs may include the full coding sequence for the mature protein (*i.e.* the protein generated when the signal polypeptide is cleaved off). In addition, the extended cDNAs obtained using the 5' ESTs may include regulatory regions upstream of the translation start site or downstream of the stop codon which control the amount, location, or developmental stage of gene expression.

As discussed above, secreted proteins are therapeutically important. Thus, the proteins expressed from the extended cDNAs or full length cDNAs obtained using the 5' ESTs may be useful in treating or controlling a variety of human conditions.

One aspect of the present invention is a purified or isolated nucleic acid having the sequence of one of SEQ ID NOs: 38-270 or having a sequence complementary thereto. In one embodiment, the nucleic acid is recombinant.

Another aspect of the present invention is a purified or isolated nucleic acid comprising at least 10 consecutive bases of the sequence of one of SEQ ID NOs: 38-270 or one of the sequences complementary thereto.

Yet another aspect of the present invention is a purified or isolated nucleic acid comprising at least 15 consecutive bases of one of the sequences of SEQ ID NOs: 38-270 or one of the sequences complementary thereto. In one embodiment, the nucleic acid is recombinant.

A further aspect of the present invention is a purified or isolated nucleic acid of at least 15 bases capable of hybridizing under stringent conditions to the sequence of one of SEQ ID NOs: 38-270 or one of the sequences complementary to the sequences of SEQ ID NOs: 38-270. In one embodiment, the nucleic acid is recombinant.

Another aspect of the present invention is a purified or isolated nucleic acid encoding a human gene product, said human gene product having a sequence partially encoded by one of the sequences of SEQ ID NO: 38-270.

Still another aspect of the present invention is a method of making a cDNA encoding a human secretory protein, said human secretory protein being partially encoded by one of SEQ ID NOs 38-270, comprising the steps of contacting a collection of mRNA molecules from human cells with a primer comprising at least 15 consecutive nucleotides of a sequence complementary to one of SEQ ID NOs: 38-270; hybridizing said primer to an mRNA in said collection that encodes said protein; reverse transcribing said hybridized primer to make a first cDNA strand from said mRNA; making a second cDNA strand complementary to said first cDNA strand; and isolating the resulting cDNA encoding said protein comprising said first cDNA strand and said second cDNA strand.

Another aspect of the invention is an isolated or purified cDNA encoding a human secretory protein, said human secretory protein comprising the protein encoded by one of SEQ ID NOs 38-270 or a fragment thereof of at least 10 amino acids, said cDNA being obtainable by the method described in the preceding paragraph. In one embodiment, the

first pair of primers comprising a second primer comprising at least 15 consecutive nucleotides of one of the sequences of SEQ ID NOs 38-270 and a third primer having a sequence therein which is included within the sequence of said first primer; performing a first polymerase chain reaction with said first pair of nested primers to generate a first PCR product; contacting said first PCR product with a second pair of primers, said second pair of primers comprising a fourth primer, said fourth primer comprising at least 15 consecutive nucleotides of said sequence of one of SEQ ID NOs: 38-270 , and a fifth primer, said fourth and fifth primers being capable of hybridizing to sequences within said first PCR product; and performing a second polymerase chain reaction, thereby generating a second PCR product.

One aspect of the present invention is an isolated or purified cDNA encoding a human secretory protein, said human secretory protein comprising the protein encoded by one of SEQ ID NOs 38-270, or a fragment thereof of at least 10 amino acids, said cDNA being obtainable by the method of the preceding paragraph. In one embodiment, the cDNA comprises the full protein coding sequence partially included in one of the sequences of SEQ ID NOs: 38-270.

Another aspect of the present invention is the method described four paragraphs above in which the second cDNA strand is made by contacting said first cDNA strand with a second primer comprising at least 15 consecutive nucleotides of the sequences of SEQ ID NOs: 38-270; hybridizing said second primer to said first strand cDNA; and extending said hybridized second primer to generate said second cDNA strand.

Another aspect of the present invention is an isolated or purified cDNA encoding a human secretory protein, said human secretory protein comprising the protein partially encoded by one of SEQ ID NOs 38-270 or comprising a fragment thereof of at least 10 amino acids, said cDNA being obtainable by the method described in the preceding paragraph. In one embodiment, the cDNA comprises the full protein coding sequence partially included in of one of the sequences of SEQ ID NOs: 38-270.

Another aspect of the present invention is a method of making a protein comprising one of the sequences of SEQ ID NOs: 271-503, comprising the steps of obtaining a cDNA encoding the full protein sequence partially included in one of the sequences of sequence of SEQ ID NOs: 38-270; inserting said cDNA in an expression vector such that said cDNA is

CLAIMS

1. A purified or isolated nucleic acid comprising the sequence of one of SEQ ID NOs: 38-270 or comprising a sequence complementary thereto.

2. The nucleic acid of Claim 1, wherein said nucleic acid is recombinant.

3. A purified or isolated nucleic acid comprising at least 10 consecutive bases of the sequence of one of SEQ ID NOs: 38-270 or one of the sequences complementary thereto.

4. A purified or isolated nucleic acid comprising at least 15 consecutive bases of one of the sequences of SEQ ID NOs: 38-270 or one of the sequences complementary thereto.

5. The nucleic acid of Claim 4, wherein said nucleic acid is recombinant.

6. A purified or isolated nucleic acid of at least 15 bases capable of hybridizing under stringent conditions to the sequence of one of SEQ ID NOs: 38-270 or one of the sequences complementary to the sequences of SEQ ID NOs: 38-270.

7. The nucleic acid of Claim 6, wherein said nucleic acid is recombinant.

8. A purified or isolated nucleic acid encoding a human gene product, said human gene product having a sequence partially encoded by one of the sequences of SEQ ID NO: 38-270.

9. A purified or isolated nucleic acid having the sequence of one of SEQ ID NOs: 38-270 or having a sequence complementary thereto.

10. A purified or isolated nucleic acid comprising the nucleotides of one of SEQ ID NOs: 38-270 which encode a signal peptide.

11. A purified or isolated polypeptides comprising a signal peptide encoded by one of the sequences of SEQ ID NOs: 38-270.

12. A vector encoding a fusion protein comprising a polypeptide and a signal peptide, said vector comprising a first nucleic acid encoding a signal peptide encoded by one of the sequences of SEQ ID NOs: 38-270 operably linked to a second nucleic acid encoding a polypeptide.

13. A method of directing the extracellular secretion of a polypeptide or the insertion of a polypeptide into the membrane comprising the steps of

19. An isolated or purified cDNA encoding a human secretory protein, said human secretory protein comprising the protein encoded by one of SEQ ID NOs 38-270 or a fragment thereof of at least 10 amino acids, said cDNA being obtainable by the method of Claim 18.

20. The cDNA of Claim 19 wherein said cDNA comprises the full protein coding sequence partially included in one of the sequences of SEQ ID NOs: 38-270.

21. The method of Claim 18, wherein the second cDNA strand is made by:

contacting said first cDNA strand with a first pair of primers, said first pair of primers comprising a second primer comprising at least 15 consecutive nucleotides of one of the sequences of SEQ ID NOs 38-270 and a third primer having a sequence therein which is included within the sequence of said first primer;

performing a first polymerase chain reaction with said first pair of nested primers to generate a first PCR product;

contacting said first PCR product with a second pair of primers, said second pair of primers comprising a fourth primer, said fourth primer comprising at least 15 consecutive nucleotides of said sequence of one of SEQ ID NO:s 38-270 , and a fifth primer, said fourth and fifth primers being capable of hybridizing to sequences within said first PCR product; and

performing a second polymerase chain reaction, thereby generating a second PCR product.

22. An isolated or purified cDNA encoding a human secretory protein, said human secretory protein comprising the protein encoded by one of SEQ ID NOs 38-270, or a fragment thereof of at least 10 amino acids, said cDNA being obtainable by the method of Claim 21.

23. The cDNA of Claim 22 wherein said cDNA comprises the full protein coding sequence partially included in one of the sequences of SEQ ID NOs: 38-270.

24. The method of Claim 18 wherein the second cDNA strand is made by:

contacting said first cDNA strand with a second primer comprising at least 15 consecutive nucleotides of the sequences of SEQ ID NOs: 38-270;

hybridizing said second primer to said first strand cDNA; and

extending said hybridized second primer to generate said second cDNA strand.

34. An isolated or purified protein comprising one of the sequences of SEQ ID NO: 271-503.

35. In an array of discrete ESTs or fragments thereof of at least 15 nucleotides in length, the improvement comprising inclusion in said array of at least one of the sequences of SEQ ID NOs: 38-270, or one of the sequences complementary to the sequences of SEQ ID NOs: 38-270, or a fragment thereof of at least 15 consecutive nucleotides.

36. The array of Claim 35 including therein at least two of the sequences of SEQ ID NOs: 38-270, the sequences complementary to the sequences of SEQ ID NOs: 38-270, or fragments thereof of at least 15 consecutive nucleotides.

37. The array of Claim 35 including therein at least five of the sequences of SEQ ID NOs: 38-270, the sequences complementary to the sequences of SEQ ID NOs: 38-270, or fragments thereof of at least 15 consecutive nucleotides.

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 36..98
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.4
seq GLSKLQFAPFSSA/LD

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 250:

```
AGTGGTTGCN GGAAGTTGAG CGGCGGCAAG AAATA ATG GCG GCA GCT ACG GGG      53
                                        Met Ala Ala Ala Thr Gly
                                           -20

GAT CCT GGA CTC TCT AAA CTG CAG TTT GCC CCT TTT AGT AGT GCC TTG     101
Asp Pro Gly Leu Ser Lys Leu Gln Phe Ala Pro Phe Ser Ser Ala Leu
-15                 -10                  -5                   1

GAT GTT GGG TTT TGG CAT GAG TTG ACC CAG AAG AAG CTG AAC GAG TAT     149
Asp Val Gly Phe Trp His Glu Leu Thr Gln Lys Lys Leu Asn Glu Tyr
             5                  10                  15

CGG CTG GAT GAA GCT CCC AAG GAC ATT AAG GGT TAT TAC TAC AAT GGT     197
Arg Leu Asp Glu Ala Pro Lys Asp Ile Lys Gly Tyr Tyr Tyr Asn Gly
         20                  25                  30

GAC TCT GCT GGG MTG CCA GCT CGC TTA ACA TTG GAG TTC AGT GCT TTT     245
Asp Ser Ala Gly Xaa Pro Ala Arg Leu Thr Leu Glu Phe Ser Ala Phe
     35                  40                  45

GAC ATG AGT GCT CCC ACC CCA AGC                                    269
Asp Met Ser Ala Pro Thr Pro Ser
 50                  55
```

(2) INFORMATION FOR SEQ ID NO: 251:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 145 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Spleen

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 39..143
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 50..154
id R50695
est

(ix) FEATURE:

(A) NAME/KEY: other
(B) LOCATION: 3..45
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 15..57
id R50695
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 81..143
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 104..166
id R94786
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 81..143
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 105..167
id T98442
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 50..130
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.4
seq LSKSLLLVPSXLS/LL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 251:

```
AAGCTTCCCC TCCCCCGGCG CCCTCTGGGG CTCCGAGCCC GGCGGGACC ATG TTC ACC   58
                                                      Met Phe Thr
                                                         -25

AGC ACC GGC TCC AGT GGG CTC TAC AAG GCG CCT CTG TCG AAG AGC CTT    106
Ser Thr Gly Ser Ser Gly Leu Tyr Lys Ala Pro Leu Ser Lys Ser Leu
            -20                 -15                 -10

CTG CTG GTC CCC AGT RCC CTC TCC CTC CTG CSC GCC CAG               145
Leu Leu Val Pro Ser Xaa Leu Ser Leu Leu Xaa Ala Gln
         -5                  1               5
```

(2) INFORMATION FOR SEQ ID NO: 252:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 427 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:

(F) TISSUE TYPE: Spleen

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -27..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.4
seq LSKSLLLVPSXLS/LL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 484:

```
Met Phe Thr Ser Thr Gly Ser Ser Gly Leu Tyr Lys Ala Pro Leu Ser
      -25                 -20                 -15

Lys Ser Leu Leu Leu Val Pro Ser Xaa Leu Ser Leu Leu Xaa Ala Gln
   -10                 -5                  1                   5
```

(2) INFORMATION FOR SEQ ID NO: 485:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 83 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Testis

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -40..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.3
seq ITLVSAAPGKVIC/EM

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 485:

```
Met Thr Ser Met Thr Gln Ser Leu Arg Glu Val Ile Lys Ala Met Thr
-40                 -35                 -30                 -25

Lys Ala Arg Asn Phe Glu Arg Val Leu Gly Lys Ile Thr Leu Val Ser
               -20                 -15                 -10

Ala Ala Pro Gly Lys Val Ile Cys Glu Met Lys Val Glu Glu Glu His
            -5                  1                   5

Thr Asn Ala Ile Gly Thr Leu His Gly Gly Leu Thr Ala Thr Leu Val
     10                  15                  20

Asp Asn Ile Ser Thr Met Ala Leu Leu Cys Thr Glu Arg Gly Ala Pro
 25                  30                  35                  40

Gly Val Ser
```

```
ID   HS695112   standard RNA; EST; 445 BP.
XX
AC   R50695;
XX
NI   g812597
XX
DT   24-MAY-1995 (Rel. 43, Created)
DT   24-MAY-1995 (Rel. 43, Last updated, Version 1)
XX
DE   yj60f12.r1 Homo sapiens cDNA clone 153167 5'.
XX
KW   EST.
XX
OS   Homo sapiens (human)
OC   Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates;
OC   Catarrhini; Hominidae; Homo.
XX
RN   [1]
RP   1-445
RA   Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M., Holman M.,
RA   Hultman M., Kucaba T., Le M., Lennon G., Marra M., Parsons J.,
RA   Rifkin L., Rohlfing T., Soares M., Tan F., Trevaskis E., Waterston R.,
RA   Williamson A., Wohldmann P., Wilson R.;
RT   "The WashU-Merck EST Project";
RL   Unpublished.
XX
CC   Contact: Wilson RK WashU-Merck EST Project Washington University
CC   School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
CC   MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:
CC   est@watson.wustl.edu High quality sequence stops: 309 Source: IMAGE
CC   Consortium, LLNL This clone is available royalty-free through LLNL
CC   ; contact the IMAGE Consortium (info@image.llnl.gov) for further
CC   information. NCBI gi: 812597
XX
FH   Key             Location/Qualifiers
FH
FT   source          1. .445
FT                   /organism="Homo sapiens"
FT                   /db_xref="ESTLIB:278"
FT                   /note="human"
FT                   /clone="153167"
XX
SQ   Sequence 445 BP; 88 A; 112 C; 103 G; 137 T; 5 other;

Hs695112  Length: 445  March 24, 1999 11:02  Type: N  Check: 6680  ..

      1  AATTCGGCAC TAGGGCTTCC CCTCCCCCGG CGTCCTCTGG GGCTCCGAGC

     51  CGGCGGGACC ATGTTCACCA GCACCGGCTC CAGTGGGCTC TACAAGGCGC

    101  CTCTGTCGAA GAGCCTTCTG CTGGTCCCCA GTGCCATCTC CCTACTGCTC

    151  GCCCTCCTCC TGCCTCACTG CCAGAAGCTC TTTGTGTATG ACCTTCACGC

    201  AGTCAAGAAC GACTTCCAGA TTTNGAGGTT GATATGTGGA AGAATAATTT
```

```
Sa218771_0040.Dna
Emest14:Hs695112

ID   HS695112   standard; RNA; EST; 445 BP.
AC   R50695;
NI   g812597
DT   24-MAY-1995 (Rel. 43, Created)
DT   24-MAY-1995 (Rel. 43, Last updated, Version 1)
DE   yj60f12.r1 Homo sapiens cDNA clone 153167 5'.
KW   EST.
OS   Homo sapiens (human)
OC   Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates;

OC   Catarrhini; Hominidae; Homo.
RN   [1]
RP   1-445
RA   Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M., Holman M.,
RA   Hultman M., Kucaba T., Le M., Lennon G., Marra M., Parsons J.,
RA   Rifkin L., Rohlfing T., Soares M., Tan F., Trevaskis E., Waterston R.,
RA   Williamson A., Wohldmann P., Wilson R.;
RT   "The WashU-Merck EST Project";
RL   Unpublished.
CC   Contact: Wilson RK WashU-Merck EST Project Washington University
CC   School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
CC   MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:
CC   est@watson.wustl.edu High quality sequence stops: 309 Source: IMAGE

CC   Consortium, LLNL This clone is available royalty-free through LLNL
CC   ; contact the IMAGE Consortium (info@image.llnl.gov) for further
CC   information. NCBI gi: 812597
FH   Key             Location/Qualifiers
FH
FT   source          1. .445
FT                   /organism="Homo sapiens"
FT                   /db_xref="ESTLIB:278"
FT                   /note="human"
FT                   /clone="153167"
SQ   Sequence 445 BP; 88 A; 112 C; 103 G; 137 T; 5 other;


SCORES    Init1: 1359  Initn: 1600  Opt: 1721 z-score: 1373.1 E():      0
  . .4% identity in 391 bp overlap

                                                    10        20        30
a218771_004                                         GGGACCATGTTCACCAGCACCGGCTCCAGT
                                                    ||||||||||||||||||||||||||||||
s695112      CCCCGGCGTCCTCTGGGGCTCCGAGCCGGCGGGACCATGTTCACCAGCACCGGCTCCAGT
                 30        40        50        60        70        80

                      40        50        60        70        80        90
a218771_004 GGGCTCTACAAGGCGCCTCTGTCGAAGAGCCTTCTGCTGGTCCCCAGTGCCCTCTCCCTC
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
s695112     GGGCTCTACAAGGCGCCTCTGTCGAAGAGCCTTCTGCTGGTCCCCAGTGCCATCTCCCTA
                90        100       110       120       130       140

                    100       110       120       130       140       150
a218771_004 CTGCTCGCCCTCCTCCTGCCTCACTGCCAGAAGCCCTTTGTGTATGACCTTCACGCAGTC
```